



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/751,574	12/29/2000	Cary Lee Bates	IBM/ 166	8366

7590

03/04/2004

Scott A. Stinebruner
Wood, Herron & Evans, L.L.P.
2700 Carew Tower
441 Vine Street
Cincinnati, OH 45202-2917

EXAMINER

HUYNH, CONG LAC T

ART UNIT

PAPER NUMBER

2178

DATE MAILED: 03/04/2004

2

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/751,574

Applicant(s)

BATES ET AL.

Examiner

Cong-Lac Huynh

Art Unit

2178

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 29 December 2000.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-43 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-43 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 29 December 2000 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
 - ☐ Certified copies of the priority documents have been received in Application No. _____.
 - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. This action is responsive to communications: the application filed on 12/29/00.
2. Claims 1-43 are pending in the case. Claims 1, 23-24, 37, 39-40, 42 are independent claims.

Claim Rejections - 35 USC § 101

3. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

4. Claims 37-39 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Specifically, independent claim 37, its dependent claim 38, and independent claim 39 state that "A program product, comprising ..." These claims are non-statutory since the program product, which includes data structures, is not claimed as embodied in computer-readable media. See MPEP 2105:

"A patent cannot be granted on such a process.

(a) Functional Descriptive Material: "Data Structures" Representing Descriptive Material Per Se or Computer Programs Representing Computer Listings Per Se

Data structures not claimed as embodied in computer-readable media are descriptive material per se and are not statutory because they are not capable of causing functional change in the computer. See, e.g., Warmerdam, 33 F.3d at 1361, 31 USPQ2d at 1760 (claim to a data structure per se held nonstatutory). Such claimed data structures do not define any structural and functional interrelationships between

Art Unit: 2178

the data structure and other claimed aspects of the invention which permit the data structure's functionality to be realized. In contrast, a claimed computer-readable medium encoded with a data structure defines structural and functional interrelationships between the data structure and the computer software and hardware components which permit the data structure's functionality to be realized, and is thus statutory."

Claim Rejections - 35 USC § 102

5. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

6. Claims 1, 9, 12, 24, 29-30, 37-38 are rejected under 35 U.S.C. 102(e) as being anticipated by Peterman (US Pat No. 6,480,838 B1, 11/12/02, filed 6/9/00, priority 4/1/98).

Regarding independent claim 1, Peterman discloses:

- scanning a plurality of documents for variants of a linguistic term (col 2, lines 40-58 and col 3, lines 51-55: database with scanned documents is searched for the words variant; col 3, lines 3-7: word variant searching of electronic documents using optical character recognition)

Art Unit: 2178

- tracking relative occurrences of a plurality of variants of the linguistic term found in the plurality of documents during scanning to determine an acceptable usage of the linguistic term (col 2, lines 52-58 and col 6, lines 25-43: determining the *possible variants of the search word* that has a probability of *occurrence* higher than a designated predetermined threshold probability *via the OCR program* inherently shows tracking the occurrences of the word variant *during scanning* to select the word variant whose occurrences *higher than a predetermined threshold* as an acceptable usage of the linguistic term)

Regarding claim 9, which is dependent on claim 1, Peterman discloses that the linguistic term comprises a single word (col 2, lines 52-58).

Regarding claim 12, which is dependent on claim 1, Peterman discloses that the plurality of variants differ from one another based upon at least one of punctuation, spelling, capitalization, hyphenation, and definition (figure 2: the alternatives for each character and combination of characters show that the variants differ from one another based upon spelling; col 10, lines 39-47).

Claims 24, 29, 30 are an apparatus of method claims 1, 9, 12, and are rejected under the same rationale.

Claims 37-38 are a program product of method claim 1, and are rejected under the same rationale.

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

9. Claims 2-8, 10, 13-14, 16-18, 20, 25-29, 31-33, 35 are rejected under 35 U.S.C. 103(a) as being unpatentable over Peterman as applied to claim 1 above, and further in view of Anick et al. (US Pat No. 6,519,586 B2, 2/11/03, filed 8/6/99).

Regarding claim 2, which is dependent on claim 1, Peterman does not disclose retrieving the plurality of documents from a network, wherein scanning the plurality of

Art Unit: 2178

documents includes scanning each document subsequent to retrieval of the document from the network.

Anick discloses that retrieving the plurality of documents from a network, wherein scanning the plurality of documents includes scanning each document subsequent to retrieval of the document from the network (col 3, lines 51-67, col 4, line 45 to col 5, line 25, col 6, lines 34-65, and figure 3A: retrieving documents from the Internet via submitting queries, then the retrieved document is searched for different terms of a word that can be used to identify a key concept or facet where searching through a document in such a way is equivalent to scanning the document).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman since Anick discloses retrieving documents from a network where scanning documents is subsequent to retrieving providing the advantage to expand the scope of searching documents within a local database as in Peterman into searching documents over the Internet.

Regarding claim 3, which is dependent on claim 2, Peterman does not disclose retrieving the plurality of documents from at least one Internet web site in response to a user browsing the at least one Internet web site, and wherein scanning the plurality of documents includes scanning each document upon retrieval of that document from the at least one Internet web site.

Anick discloses retrieving the plurality of documents from at least one Internet web site in response to a user browsing the at least one Internet web site, and wherein scanning

the plurality of documents includes scanning each document upon retrieval of that document from the at least one Internet web site (col 3, lines 51-67, figure 3A: retrieving documents from a Internet web site in response to user entering the topic, then the searching (equivalent to scanning as mentioned in claim 2) is performed on the retrieved document from said web site).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman since Anick discloses retrieving documents from a network where scanning documents is subsequent to retrieving providing the advantage to expand the scope of searching documents within a local database as in Peterman into searching documents over the Internet.

Regarding claim 4, which is dependent on claim 2, Peterman and Anick do not disclose determining whether a retrieved document has already been scanned before scanning the retrieved document.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Peteman and Anick to incorporate determining whether a retrieved document has already been scanned before scanning the retrieved document for the following reason. It was well known that each computer system has a cache for storing the retrieved documents as well as a cookie file for storing user actions performed while in the Internet and for preventing a repeated action if it is already in cache. Therefore, the system will check to see if a document is scanned yet via the cache and cookie file before performing scanning documents.

Regarding claim 5, which is dependent on claim 2, Peterman does not disclose determining whether to scan a retrieved document based upon a source parameter associated with the linguistic term.

Anick discloses determining whether to scan a retrieved document based upon a source parameter associated with the linguistic term (col 8, line 18 to col 9, line 67: a list of parameters provided to users for changing the selection of facets derived from the retrieved documents).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick to Peterman since Anick determining whether to scan a retrieved document based upon a source parameter providing the advantage of selecting different options on scanning via selecting an appropriate parameter.

Regarding claim 6, which is dependent on claim 1, Peterman does not disclose browsing a second plurality of documents retrieved from at least one Internet web site in response to user input, wherein scanning the first plurality of documents is performed concurrently with browsing the second plurality of documents.

Anick discloses:

- browsing a plurality of documents retrieved from at least one Internet web site in response to user input (col 3, lines 51-58; figures 3A-B: searching and browsing web pages in response to user input of topics on the Internet)
- scanning a plurality of documents on the Internet (col 5, lines 52-63)

Anick does not explicitly disclose that the above two steps occur concurrently.

Art Unit: 2178

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to include the concurrency of said two steps into Anick for the following reason. Anick discloses that browsing for a desired document on the Internet and measuring the facet frequency are occurred on a web page (figures 3A-B; col 5, lines 52-63 and col 6, lines 34-65). The fact that the web browser receives the user input on the frames of a web page for performing the two different functions suggests that the two steps can be carried out concurrently to preserve the consistence of data in the frames of the web page.

Regarding claim 7, which is dependent on claim 1, Peterman does not disclose that scanning the first plurality of documents is performed in a background thread while documents from the second plurality of documents are being browsed.

As mentioned in claim 6 above, Anick suggests that steps scanning and browsing documents be occurred concurrently for preserving the consistency of data presented in the web page. Therefore, it is further suggested that the two actions, scanning documents and browsing documents be performed parallelly in two different threads.

Regarding claim 8, which is dependent on claim 7, Peterman does not disclose that scanning the first plurality of documents includes scanning documents stored in a local history cache. However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Peterman to incorporate scanning documents stored in a local history cache since it was well known that cache is a local

storage of a computer with a plurality of stored documents. Therefore, scanning documents includes documents stored in cache.

Regarding claim 10, which is dependent on claim 1, Peterman does not disclose that the linguistic term comprises a phrase.

Anick discloses that the linguistic term comprises a phrase (figure 3B and col 5, lines 52-63).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for expanding the scope of considering the variant of a linguistic term as a phrase including at least two words instead of a single word as in Peterman.

Regarding claim 13, which is dependent on claim 1, Peterman does not disclose that scanning the plurality of documents includes scanning a document for an enumerated variant of the linguistic term.

Anick discloses scanning the plurality of documents includes scanning a document for an enumerated variant of the linguistic term (col 6, lines 20-33: the fact that the facet list is retained for successive searches shows that the number of facets, which are the variants of a linguistic term, is counted and determined; in other words, the scanning is for an enumerated variant of a linguistic term).

Art Unit: 2178

Regarding claim 14, which is dependent on claim 1, Peterman does not disclose that scanning the plurality of documents includes scanning a document for an unenumerated variant of the linguistic term.

Anick discloses that scanning the plurality of documents includes scanning a document for an unenumerated variant of the linguistic term (col 5, lines 52-63: the fact that a user can choose *to list any desired number of facet phrases* in response to the user's click of a facet selection box for the facets corresponding to the scanned documents suggests an unenumerated variant of the linguistic term when scanning a document).

Regarding claim 16, which is dependent on claim 1, Peterman does not disclose that tracking relative occurrences of the plurality of variants includes weighting occurrences based upon locations of such occurrences within the plurality of documents.

Anick discloses tracking relative occurrences of the plurality of variants includes weighting occurrences based upon locations of such occurrences within the plurality of documents (col 10, 1-24: weighting facet occurrences on only the terms *appearing near the beginning of the documents*, those appearing in the first n document lines).

Regarding claim 17, which is dependent on claim 1, Peterman does not disclose that tracking relative occurrences of the plurality of variants includes weighting occurrences based upon document types of the documents within which such occurrences are found.

Anick discloses that tracking relative occurrences of the plurality of variants includes weighting occurrences based upon document types of the documents within which such occurrences are found (figure 3B, col 10, lines 12-24, col 3, lines 51-67: weighting the density of the terms, which is equivalent to the term occurrences, in a document where a document can be a form of articles, papers, statements, correspondence, etc.).

Regarding claim 18, which is dependent on claim 1, Peterman does not disclose storing a variant of the linguistic term in an electronic dictionary.

Anick discloses storing a variant of the linguistic term in an electronic dictionary (figures 3A-B: lists of variants of linguistic terms are stored and displayed for users to select).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for the following reason. Anick discloses storing lists of variants of linguistic terms for users to select providing the advantage to apply said storing to Peterman to store the variants of linguistic terms with high occurrences for rapidly searching and retrieving documents.

Regarding claim 20, which is dependent on claim 1, Peterman does not disclose tracking relative occurrences of the plurality of variants includes storing context information associated with each occurrence of a variant of the linguistic term.

Anick discloses tracking relative occurrences of the plurality of variants includes storing context information associated with each occurrence of a variant of the linguistic term (col 6, lines 46-65: encoding strong, long-live information relating to relationships and

Art Unit: 2178

attributes which serve well as the lexical constructs from which dispersion rates are derived suggests that said information relating to the occurrences of the variants for a particular word be stored in a database for searching or retrieving documents where the database of words and phrases is considered as similar as an electronic dictionary for storing terms).

Claims 25-28 are for an apparatus of method claims 2, 4, 5, 7, and are rejected under the same rationale.

Claims 29, 31-33, 35 are for an apparatus of method claims 10, 13-14, 16, 20 and are rejected under the same rationale.

10. Claim 11 is rejected under 35 U.S.C. 103(a) as being unpatentable over Peterman as applied to claim 1 above, and further in view of Sundaresan et al. (US Pat No. 6,385,629 B1, 5/7/02, filed 11/15/99).

Regarding claim 11, which is dependent on claim 1, Peterman does not disclose that the linguistic term comprises an acronym.

Sundaresan discloses scanning web documents for potential (acronym, expansion) pairs as well as identifying their occurrences (col 5, lines 9-16; col 2, line 64 to col 3, line 7).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Sundaresan into Peterman since Sundaresan discloses

Art Unit: 2178

scanning web documents for potential (acronym, expansion) pairs as well as identifying their occurrences providing the advantage to incorporate Sundaresan feature into Peterman wherein a linguistic term being a word since *the text of web documents* in Sundaresan *containing acronyms* shows that the *linguistic term can be an acronym*.

11. Claim 15 is rejected under 35 U.S.C. 103(a) as being unpatentable over Peterman and Anick as applied to claims 1 and 14 above, and further in view of Komissarchik et al. (US Pat No. 5,799,276, 8/25/98, filed 11/7/95).

Regarding claim 15, which is dependent on claim 14, Peterman and Anick do not disclose that scanning the document for the unenumerated variant of the linguistic term includes scanning the document using phonetic comparison.

Komissarchik discloses that a single phonetic transcription of a word in a dictionary can be replaced with *a plurality of phonetic transactions of phonetically permissible variants of the word* (col 3, line 63 to col 4, line 15 and col 11, line 50 to col 12, line 13).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Komissarchik into Peterman and Anick since Komissarchik teaches that it is possible to replace a word in a dictionary with phonetically permissible variants of the word thus motivating to phonetically scanning documents for the variants of a word instead of lexically scanning as in Peterman and Anick utilizing the phonetic comparison.

12. Claims 19, 21-22, 34, 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Peterman and Anick as applied to claims 1 and 18 above, and further in view of Mueller (US Pat No. 6,694,484 B1, 2/17/04, filed 6/3/97).

Regarding claim 19, which is dependent on claim 18, Peterman does not disclose spell checking a document using the electronic dictionary subsequent to storing the variant in the electronic dictionary.

Mueller discloses spell checking on HTML documents after the HTML documents are stored in the database (col 8, lines 9-35).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Mueller into Peterman for the following reason. Mueller discloses spell checking the HTML document using the spell checker program, where it was well known that any spell checker program uses an electronic dictionary for checking the spelling, and where these documents are stored in a database thus motivating to apply the spell checking program on a document subsequent to storing the variant in the electronic dictionary in Peterman since a whole document when being stored includes a plurality of words which can be the variant of a linguistic term.

Regarding claim 21, which is dependent on claim 1, Peterman does not disclose scanning a document for a spell definition tag that identifies a variant of the linguistic term.

Art Unit: 2178

Mueller discloses scanning a document for a spell definition tag that identifies a variant of the linguistic term (col 3, lines 25-35 and col 8, lines 9-35: each application program has an association tag, and so the spell check program also has an association tags embedded in the HTML document for identifying the spell checker program).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Mueller into Peterman since Mueller teaches scanning a document for a spell definition tag embedded in the HTML document for performing the spell checking providing the advantage to apply the spell check feature into Peterman documents for rapidly and effectively providing users documents without spell errors.

Regarding claim 22, which is dependent on claim 1, Peterman and Mueller do not disclose scanning the plurality of documents and tracking relative occurrences are performed responsive to detecting a variant of the linguistic term during spell checking of a document.

Instead, Peterman does teach scanning the plurality of documents and tracking relative occurrences of the variant of a linguistic term where tracking occurrences of the variant of a linguistic term inherently shows detecting a variant of the linguistic term (col 2, lines 40-58, col 3, lines 51-55, col 6, lines 25-43).

Mueller discloses scanning a document for spell checking of the document (col 8, lines 9-35).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Mueller into Peterman to incorporate the spell checking

program of Mueller into scanning documents of Peterman for detecting the variant of a linguistic term since scanning a document in Peterman is for looking through the document for detecting the variant and scanning a document in Mueller is also for looking through the document for detecting spell errors.

Claims 34 and 36 are for an apparatus of method claims 19 and 21, and are rejected under the same rationale.

13. Claims 23 and 39 are rejected under 35 U.S.C. 103(a) as being unpatentable over Anick et al. (US Pat No. 6,519,586, 2/11/03, filed 8/6/99).

Regarding independent claim 23, Anick discloses:

- browsing a plurality of web sites on the Internet in response to user input (col 3, lines 51-58; figures 3A-B: searching and browsing web pages in response to user input of topics on the Internet)
- tracking relative occurrences of a plurality of variants of a linguistic term found in the plurality of web sites to determine an acceptable usage of the linguistic term (col 5, lines 52-63: measuring the frequency of facet's occurrence for a retrieved document)

Anick does not explicitly disclose that the above two steps occur concurrently.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to include the concurrency of said two steps into Anick for the following reason. Anick discloses that browsing for a desired

Art Unit: 2178

document on the Internet and measuring the facet frequency are occurred on a web page (figures 3A-B; col 5, lines 52-63 and col 6, lines 34-65). The fact that the web browser receives the user inputs on the frames of a web page for performing the two different functions suggests that the two steps can be carried out concurrently to preserve the consistence of data on the web page.

Claim 39 is for a program product of method claim 23, and is rejected under the same rationale.

14. Claims 40-43 are rejected under 35 U.S.C. 103(a) as being unpatentable over Mueller (US Pat No. 6,694,484 B1, 2/17/04, filed 6/3/97) in view of Anick et al. (US Pat No. 6,519,586, 2/11/03, filed 8/6/99).

Regarding independent claim 40, Mueller discloses:

- comparing terms in a first document against an electronic dictionary (col 8, lines 9-35: the spell checker program *scans through the HTML document confirming that each word constitutes a properly spelled word* where it was well known that the spell check program has an associated electronic dictionary to perform comparing for confirming word spelling)
- storing data identifying the misspelled word and its locations in the HTML document (col 8, lines 18-25)

Mueller does not disclose:

- in response to determining during the comparison that a term from the document is not in the electronic dictionary, automatically scanning a plurality of documents from the Internet to identify at least one acceptable usage of the term

Anick discloses scanning a plurality of documents from the Internet to identify at least one acceptable usage of a term in a document (col 5, lines 52-63).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Mueller since Anick teaches scanning a plurality of documents from the Internet to identify at least one acceptable usage of a term in a document thus motivating to scan Internet documents to identify at least one acceptable usage of *any term in any document* where said term is not found in an electronic dictionary or found misspelled in the spell checking of Mueller.

Regarding claim 41, which is dependent on claim 40, Anick discloses:

- tracking relative occurrences of a plurality of variants of the term found in the plurality of documents (col 5, lines 52-63: measuring the frequency of facet's occurrence; col 10, lines 1-24: the term density in a document shows the occurrences of a term in a document)
- displaying results of such tracking to a user (figure 3B and col 5, lines 52-63)

Regarding independent claim 42, Mueller discloses:

- detecting a tag for spell check application within a document retrieved from the Internet that identifies an acceptable variant of a linguistic term for carrying out

spell checking (col 3, lines 52-60 and col 8, lines 9-35: the association tag embedded in the HTML document retrieved from the Internet associated with the spell checker program where the terms that do not constitutes a properly spelled word are identified and prompted to user with several suggested replacements; further, it was well known that a spell check program further allows users to enter a word to replace the incorrect word which is not included in the suggested replacement or the electronic dictionary provided by the system)

- in response to detecting said tag, automatically adding the acceptable variant of the linguistic term to an electronic dictionary (as mentioned in the detecting step, the word entered by users will be automatically added to the electronic dictionary of the spell check program)

Mueller does not explicitly disclose such a spell definition tag. However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Mueller to include the spell definition tag since Mueller discloses association tags associated with application programs including the spell checking program thus motivating to apply the association tag with specific information as a spell definition tag for performing the spell checking.

Regarding claim 43, which is dependent on claim 42, Mueller discloses that detecting the spell definition tag is performed during user browsing of the Internet (col 8, lines 9-35; figure 4: searching for an association tag in an *received HTML document* from a computer system *over an external communication link* where the association tag related

Art Unit: 2178

to the spell check program (as mentioned in claim 42) indicates that detecting the spell check tag is performed during user browsing the Internet).

15. Claims 1-10, 12-18, 20, 23-33, 35, 37-39 are rejected under 35 U.S.C. 103(a) as being unpatentable over Anick et al. (US Pat No. 6,519,586 B2, 2/11/03, filed 8/6/99).

Regarding independent claim 1, Anick discloses:

- searching a plurality of documents for variants of a linguistic term (figures 3A-B: searching for variants of a linguistic term in the retrieved documents)
- tracking relative occurrences of a plurality of variants of the linguistic term found in the plurality of documents during searching to determine an acceptable usage of the linguistic term (col 5, lines 52-63: measuring the frequency of facet's occurrences in the documents)

Anick does not use the word "scanning document" for variants of a linguistic term.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to have scanning documents for variant of a linguistic term since searching said variant in the documents means looking through the documents to find out the word variant. That is scanning the documents for the word variant.

Regarding claim 2, which is dependent on claim 1, Anick discloses that retrieving the plurality of documents from a network, wherein scanning the plurality of documents includes scanning each document subsequent to retrieval of the document from the

Art Unit: 2178

network (col 3, lines 51-67, col 4, line 45 to col 5, line 25, col 6, lines 34-65, and figure 3A: retrieving documents from the Internet via submitting queries, then the retrieved document is searched for different terms of a word that can be used to identify a key concept or facet where searching through a document in such a way is equivalent to scanning the document).

Regarding claim 3, which is dependent on claim 2, Anick discloses retrieving the plurality of documents from at least one Internet web site in response to a user browsing the at least one Internet web site, and wherein scanning the plurality of documents includes scanning each document upon retrieval of that document from the at least one Internet web site (col 3, lines 51-67, figure 3A: retrieving documents from a Internet web site in response to user entering the topic, then the searching (equivalent to scanning as mentioned in claim 2) is performed on the retrieved document from said web site).

Regarding claim 4, which is dependent on claim 2, Anick does not disclose determining whether a retrieved document has already been scanned before scanning the retrieved document.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to incorporate determining whether a retrieved document has already been scanned before scanning the retrieved document for the following reason. It was well known that each computer system has a cache for storing the retrieved documents as well as a cookie file for storing user actions

performed while in the Internet and for preventing a repeated action if it is already in cache. Therefore, the system will check to see if a document is scanned yet via the cache and cookie file before performing scanning documents.

Regarding claim 5, which is dependent on claim 2, Anick discloses determining whether to scan a retrieved document based upon a source parameter associated with the linguistic term (col 8, line 18 to col 9, line 67: a list of parameters provided to users for changing the selection of facets derived from the retrieved documents).

Regarding claim 6, which is dependent on claim 1, Anick discloses:

- browsing a plurality of documents retrieved from at least one Internet web site in response to user input (col 3, lines 51-58; figures 3A-B: searching and browsing web pages in response to user input of topics on the Internet)
- scanning a plurality of documents on the Internet (col 5, lines 52-63)

Anick does not explicitly disclose that the above two steps occur concurrently.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to include the concurrency of said two steps into Anick for the following reason. Anick discloses that browsing for a desired document on the Internet and measuring the facet frequency are occurred on a web page (figures 3A-B; col 5, lines 52-63 and col 6, lines 34-65). The fact that the web browser receives the user input on the frames of a web page for performing the two

different functions suggests that the two steps can be carried out concurrently to preserve the consistence of data in the frames of the web page.

Regarding claim 7, which is dependent on claim 1, as mentioned in claim 6 above, Anick suggests that steps scanning and browsing documents be occurred concurrently for preserving the consistency of data presented in the web page. Therefore, it is further suggested that the two actions, scanning documents and browsing documents be performed parallely in two different threads.

Regarding claim 8, which is dependent on claim 7, Anick does not disclose that scanning the first plurality of documents includes scanning documents stored in a local history cache. However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Peterman to incorporate scanning documents stored in a local history cache since it was well known that cache is a local storage of a computer with a plurality of stored documents. Therefore, scanning documents includes documents stored in cache.

Regarding claim 9, which is dependent on claim 1, Anick does not disclose that the linguistic term comprises a word. Instead, Anick discloses that the linguistic term comprises a phrase (figure 3B and col 5, lines 52-63).

Art Unit: 2178

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to include that the linguistic term comprises a word since the phrase in Anick includes at least a word.

Regarding claim 10, which is dependent on claim 1, Anick discloses that the linguistic term comprises a phrase (figure 3B and col 5, lines 52-63).

Regarding claim 12, which is dependent on claim 1, Anick discloses that the plurality of variants differ from one another based upon at least one of punctuation, spelling, capitalization, hyphenation, and definition (figure 3B: variants in 325 differ from definition).

Regarding claim 13, which is dependent on claim 1, Anick discloses scanning the plurality of documents includes scanning a document for an enumerated variant of the linguistic term (col 6, lines 20-33: the fact that the facet list is retained for successive searches shows that the number of facets, which are the variants of a linguistic term, is counted and determined; in other words, the scanning is for an enumerated variant of a linguistic term).

Regarding claim 14, which is dependent on claim 1, Anick discloses that scanning the plurality of documents includes scanning a document for an unenumerated variant of the linguistic term (col 5, lines 52-63: the fact that a user can choose *to list any desired*

Art Unit: 2178

number of facet phrases in response to the user's click of a facet selection box for the facets corresponding to the scanned documents suggests an unenumerated variant of the linguistic term when scanning a document).

Regarding claim 16, which is dependent on claim 1, Anick discloses tracking relative occurrences of the plurality of variants includes weighting occurrences based upon locations of such occurrences within the plurality of documents (col 10, 1-24: weighting facet occurrences on only the terms *appearing near the beginning of the documents*, those appearing in the first n document lines).

Regarding claim 17, which is dependent on claim 1, Anick discloses that tracking relative occurrences of the plurality of variants includes weighting occurrences based upon document types of the documents within which such occurrences are found (figure 3B, col 10, lines 12-24, col 3, lines 51-67: weighting the density of the terms, which is equivalent to the term occurrences, in a document where a document can be a form of articles, papers, statements, correspondence, etc.).

Regarding claim 18, which is dependent on claim 1, Anick discloses storing a variant of the linguistic term in an electronic dictionary (figures 3A-B: lists of variants of linguistic terms are stored and displayed for users to select).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for the following reason. Anick

Art Unit: 2178

discloses storing lists of variants of linguistic terms for users to select providing the advantage to apply said storing to Peterman to store the variants of linguistic terms with high occurrences for rapidly searching and retrieving documents.

Regarding claim 20, which is dependent on claim 1, Anick discloses tracking relative occurrences of the plurality of variants includes storing context information associated with each occurrence of a variant of the linguistic term (col 6, lines 46-65: encoding strong, long-live information relating to relationships and attributes which serve well as the lexical constructs from which dispersion rates are derived suggests that said information relating to the occurrences of the variants for a particular word be stored in a database for searching or retrieving documents where the database of words and phrases is considered as similar as an electronic dictionary for storing terms).

Regarding independent claim 23, Anick discloses:

- browsing a plurality of web sites on the Internet in response to user input (col 3, lines 51-58; figures 3A-B: searching and browsing web pages in response to user input of topics on the Internet)
- tracking relative occurrences of a plurality of variants of a linguistic term found in the plurality of web sites to determine an acceptable usage of the linguistic term (col 5, lines 52-63: measuring the frequency of facet's occurrence for a retrieved document)

Anick does not explicitly disclose that the above two steps occur concurrently.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to include the concurrency of said two steps into Anick for the following reason. Anick discloses that browsing for a desired document on the Internet and measuring the facet frequency are occurred on a web page (figures 3A-B; col 5, lines 52-63 and col 6, lines 34-65). The fact that the web browser receives the user inputs on the frames of a web page for performing the two different functions suggests that the two steps can be carried out concurrently to preserve the consistence of data on the web page.

Claims 24-29 are for an apparatus of method claims 1, 2, 4, 5, 7, and are rejected under the same rationale.

Claims 30-33, 35 are for an apparatus of method claims 12-14,16, 20 and are rejected under the same rationale.

Claims 37-39 are for a program product of method claims 1 and 23, and are rejected under the same rationale.

16. Claim 11 is rejected under 35 U.S.C. 103(a) as being unpatentable over Anick as applied to claim 1 above, and further in view of Sundaresan et al. (US Pat No. 6,385,629 B1, 5/7/02, filed 11/15/99).

Art Unit: 2178

Regarding claim 11, which is dependent on claim 1, Anick does not disclose that the linguistic term comprises an acronym.

Sundaresan discloses scanning web documents for potential (acronym, expansion) pairs as well as identifying their occurrences (col 5, lines 9-16; col 2, line 64 to col 3, line 7).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Sundaresan into Anick since Sundaresan discloses scanning web documents for potential (acronym, expansion) pairs as well as identifying their occurrences providing the advantage to incorporate Sundaresan feature into Anick wherein a linguistic term being a word since *the text of web documents* in Sundaresan *containing acronyms* shows that the *linguistic term can be an acronym*.

17. Claim 15 is rejected under 35 U.S.C. 103(a) as being unpatentable over Anick as applied to claims 1 and 14 above, and further in view of Komissarchik et al. (US Pat No. 5,799,276, 8/25/98, filed 11/7/95).

Regarding claim 15, which is dependent on claim 14, Anick does not disclose that scanning the document for the unenumerated variant of the linguistic term includes scanning the document using phonetic comparison.

Komissarchik discloses that a single phonetic transcription of a word in a dictionary can be replaced with *a plurality of phonetic transactions of phonetically permissible variants of the word* (col 3, line 63 to col 4, line 15 and col 11, line 50 to col 12, line 13).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Komissarchik into Anick since Komissarchik teaches that it is possible to replace a word in a dictionary with phonetically permissible variants of the word thus motivating to phonetically scanning documents for the variants of a word instead of lexically scanning as in Anick, utilizing the phonetic comparison.

18. Claims 19, 21-22, 34, 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Anick as applied to claims 1 and 18 above, and further in view of Mueller (US Pat No. 6,694,484 B1, 2/17/04, filed 6/3/97).

Regarding claim 19, which is dependent on claim 18, Anick does not disclose spell checking a document using the electronic dictionary subsequent to storing the variant in the electronic dictionary.

Mueller discloses spell checking on HTML documents after the HTML documents are stored in the database (col 8, lines 9-35).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Mueller into Anick for the following reason. Mueller discloses spell checking the HTML document using the spell checker program, where it was well known that any spell checker program uses an electronic dictionary for checking the spelling, and where these documents are stored in a database thus motivating to apply the spell checking program on a document subsequent to storing the

variant in the electronic dictionary in Anick since a whole document when being stored includes a plurality of words which can be the variant of a linguistic term.

Regarding claim 21, which is dependent on claim 1, Anick does not disclose scanning a document for a spell definition tag that identifies a variant of the linguistic term.

Mueller discloses scanning a document for a spell definition tag that identifies a variant of the linguistic term (col 3, lines 25-35 and col 8, lines 9-35: each application program has an association tag, and so the spell check program also has an association tags embedded in the HTML document for identifying the spell checker program).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Mueller into Anick since Mueller teaches scanning a document for a spell definition tag embedded in the HTML document for performing the spell checking providing the advantage to apply the spell check feature into Anick documents for rapidly and effectively providing users documents without spell errors.

Regarding claim 22, which is dependent on claim 1, Anick and Mueller do not disclose scanning the plurality of documents and tracking relative occurrences are performed responsive to detecting a variant of the linguistic term during spell checking of a document.

Instead, Anick does teach scanning the plurality of documents and tracking relative occurrences of the variant of a linguistic term where tracking occurrences of the variant

Art Unit: 2178

of a linguistic term inherently shows detecting a variant of the linguistic term (figures 3A-B and col 5, lines 52-63).

Mueller discloses scanning a document for spell checking of the document (col 8, lines 9-35).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Mueller into Anick to incorporate the spell checking program of Mueller into scanning documents of Anick for detecting the variant of a linguistic term since scanning a document in Anick is for looking through the document for detecting the variant and scanning a document in Mueller is also for looking through the document for detecting spell errors.

Claims 34 and 36 are for an apparatus of method claims 19 and 21, and are rejected under the same rationale.

Conclusion

19. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Irons (US Pat No. 6,192,165 B1, 2/20/01, filed 12/30/97).

Katariya et al. (US Pat No. 6,549,897 B1, 4/15/03, filed 12/17/98).

Wynn et al. (US Pat No. 6,493,000 B1, 12/10/02, filed 2/17/00).

Cullen et al. (US Pat No. 6,397,213 B1, 5/28/02, filed 5/12/99).

Cai et al. (US Pat No. 6,175,834 B1, 1/16/01, filed 6/24/98).

Art Unit: 2178

McKeown et al. (US Pat No. 6,473,730 B1, 10/29/02, filed 4/12/99).

Walfish et al. (US Pat No. 6,047,300, 4/4/00, filed 5/15/97).

Borovoy et al. (US Pat No. 5,873,107, 2/16/99, filed 3/29/96).

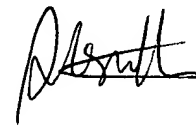
Golding et al. (US Pat No. 5,956,739, 9/21/99, filed 6/25/96).

20. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Cong-Lac Huynh whose telephone number is 703-305-0432. The examiner can normally be reached on Mon-Fri (8:30-6:00).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on 703-308-5186. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Clh
2/21/04



STEPHEN S. HONG
PRIMARY EXAMINER